Teodor Fredriksson
May 9, 2024

Supplementary Information

Bayesian Data Analysis

This section outlines how the Bayesian Data Analysis was performed. We follow the recommendations and guidlines of [1, 2, 3] amd asses the Markov Chain Monte Carlo (MCMC) chains and Posterior Predictive Checks.

Assesing the MCMC chains. All analysis were conducted using the `R` programming language and the `cmdstanr` library. Each model was computed using four parallel chains using 2000 iterations and an additional 200 warmup iterations. None of the iteration diverged as can be seen in the traceplots Figures 1, 2, 3, 4.

In addition we investigate diagnostics of the posterior estimates, such as *Gelman-Rubin Potential Scale Reduction*($\widehat{R}$) [4] the number of *efficient samples* (n_{eff}). Tables 1, 2, 3, contains three columns. The first column contains $\widehat{R}$ and the other two contains two different estimates of the number of effective samples. As a rule of thumb we should have $\widehat{R} < 1.01$, which is satisfied in most scenarios except some where where $\widehat{R}$ is slightly higher but we are happy with that. The number of efficient samples should be at least 200, which is satisfied in all cases.

Posterior Predictive Checks. For posterior predictive checks, we investigate how well the mean error rate is captured by the posterior distribution. Figures 5, shows the posterior distribution of the mean for each algorithm and the Figures show that the mean is well captured by the posterior distributions.

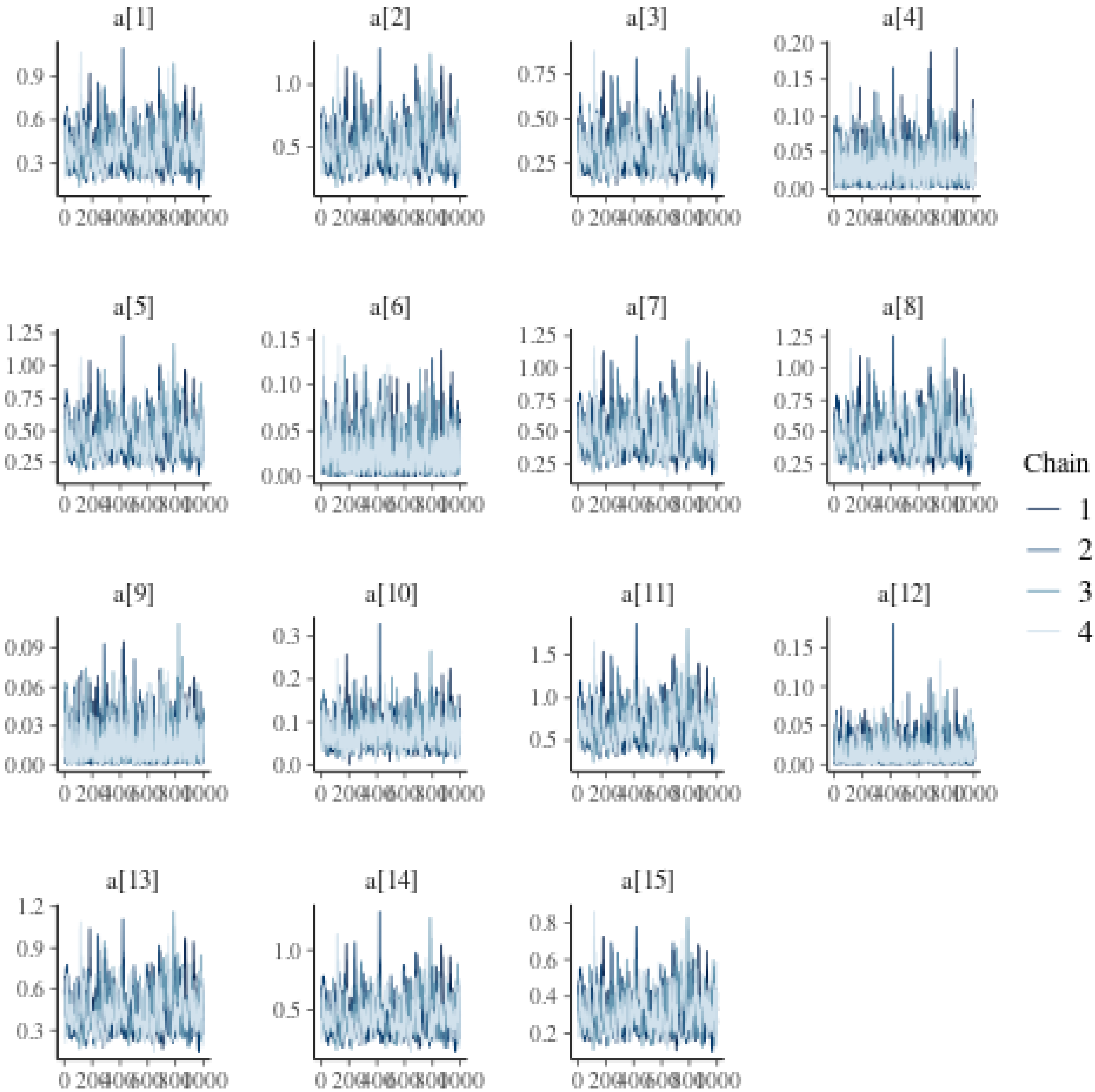

FIGURE 1. Simulated MCMC chains for the a_i parameters.

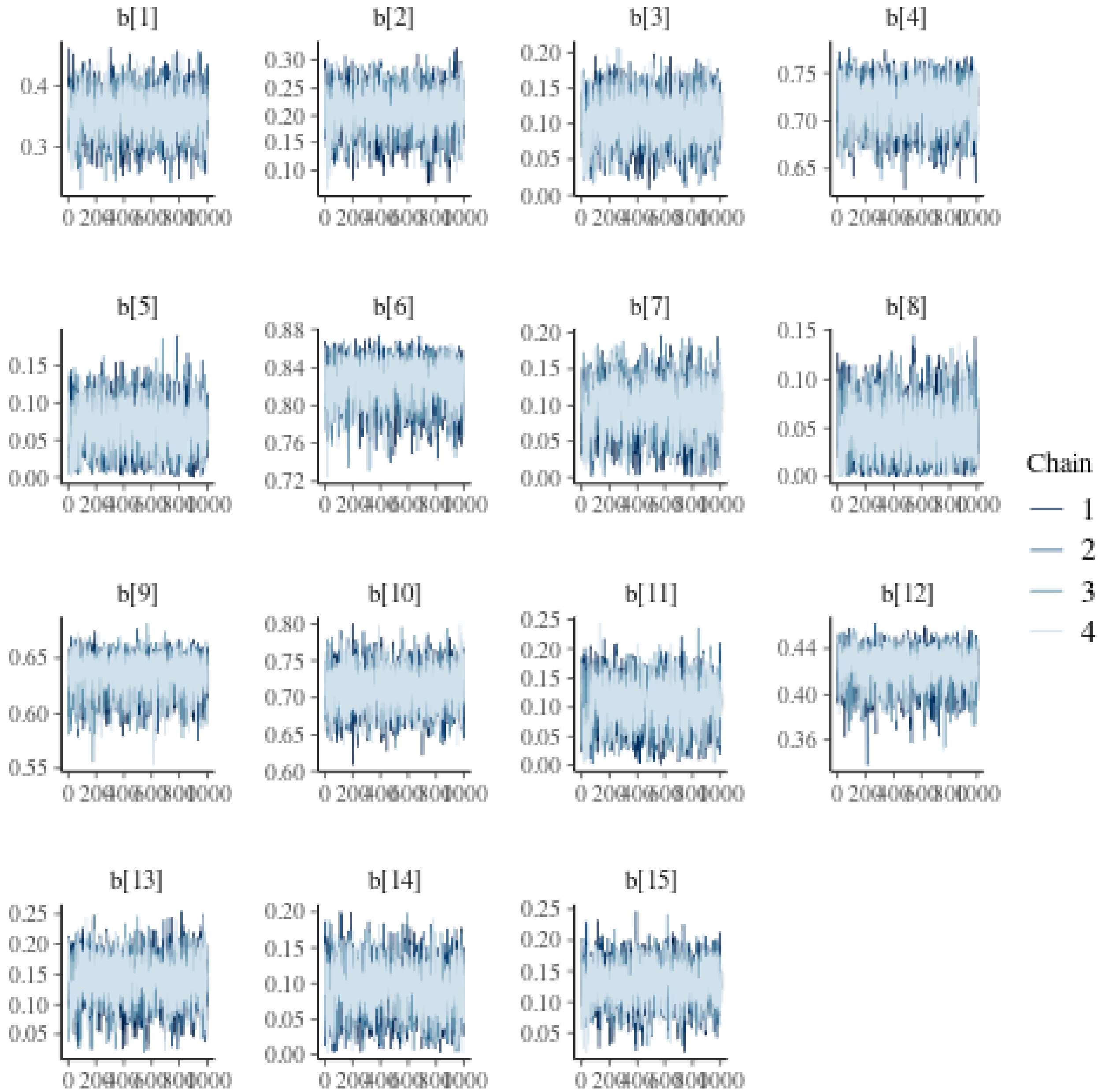

FIGURE 2. Simulated MCMC chains for the b_i parameters.

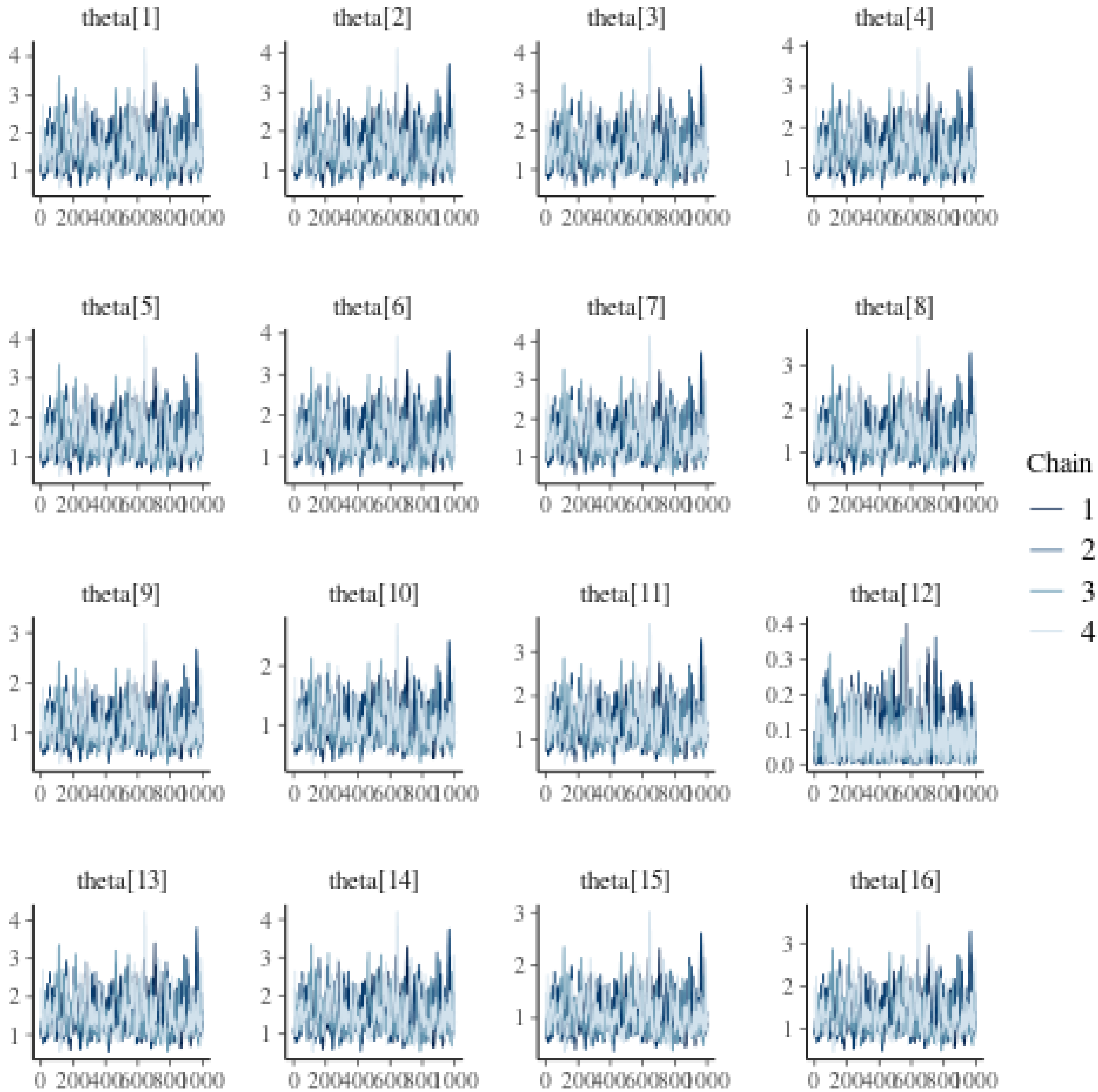

FIGURE 3. Simulated MCMC chains for the θ_i parameter

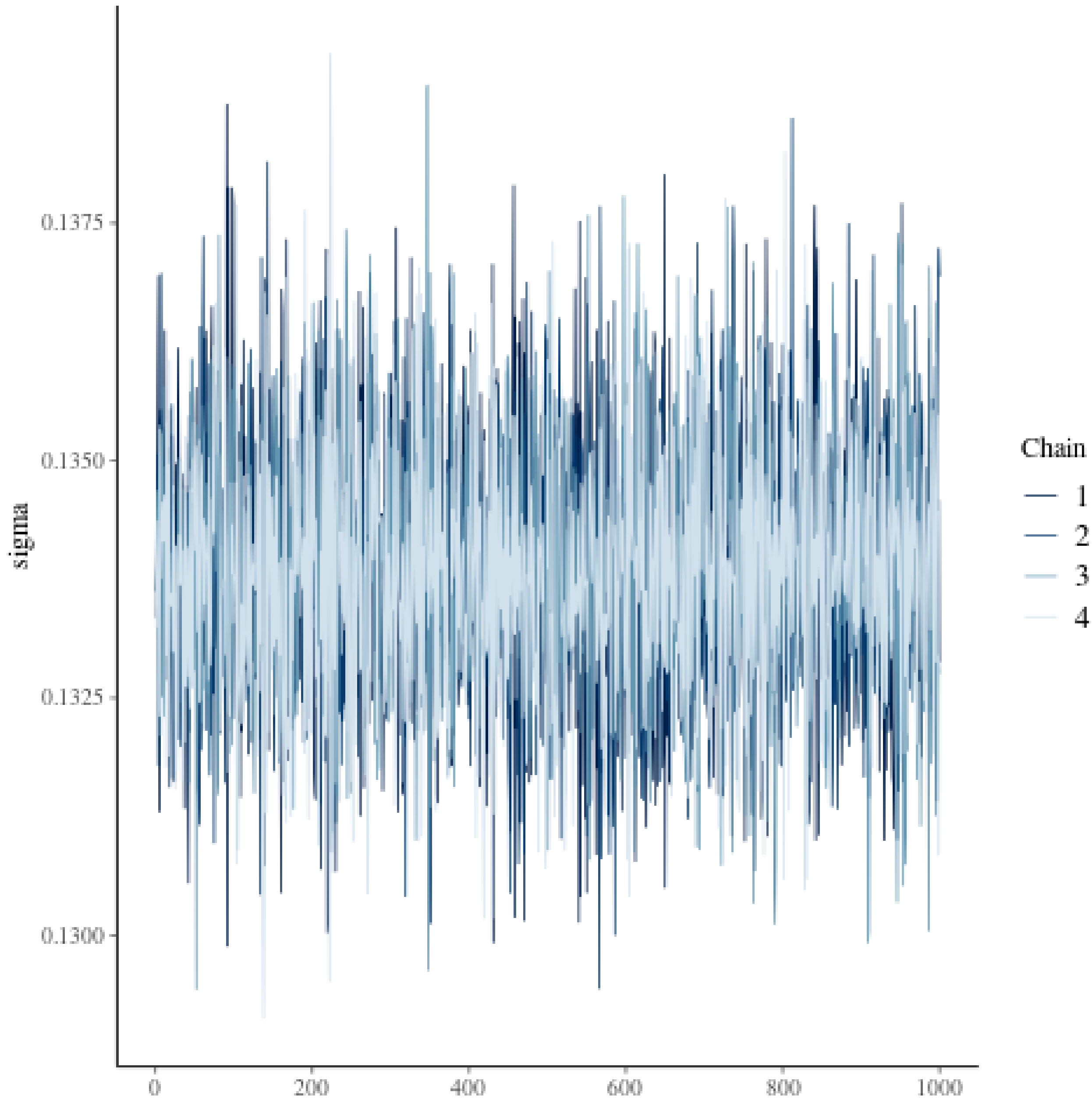

Figure 4. Simulated MCMC chains for the strength parameters σ_i

Table 1. Diagnostics for the posterior estimates of a and b.

variable	Rhat	ess_bulk	ess_tail
a[1]	1.016	409.737	751.035
a[2]	1.017	402.467	697.120
a[3]	1.018	410.370	769.547
a[4]	1.005	1058.330	1767.918
a[5]	1.017	397.421	662.110
a[6]	1.004	1123.990	1572.976
a[7]	1.016	402.374	741.374
a[8]	1.018	396.616	701.782
a[9]	1.001	1615.576	1908.252
a[10]	1.009	693.232	1385.871
a[11]	1.017	395.447	615.851
a[12]	1.003	1566.255	1987.305
a[13]	1.016	409.168	824.098
a[14]	1.018	387.667	695.829
a[15]	1.017	417.107	812.848
b[1]	1.001	1696.522	2062.004
b[2]	1.001	1427.749	1505.836
b[3]	1.001	1601.150	1124.897
b[4]	1.002	2004.280	2292.972
b[5]	1.000	1159.704	886.296
b[6]	1.002	2297.903	2463.459
b[7]	1.001	1177.618	939.356
b[8]	1.000	1417.486	1668.041
b[9]	1.001	2458.670	2581.657
b[10]	1.001	2977.649	2454.494
b[11]	1.000	881.180	877.535
b[12]	1.003	2598.330	2485.031
b[13]	1.002	1632.760	1469.293
b[14]	1.000	1381.927	1502.449
b[15]	1.001	1808.447	1642.803

Table 2. Diagnostics for the posterior estimates of θ.

variable	Rhat	ess_bulk	ess_tail
theta[1]	1.018	394.019	664.839
theta[2]	1.018	395.767	651.947
theta[3]	1.018	394.049	669.624
theta[4]	1.018	392.865	674.455
theta[5]	1.018	394.948	703.921
theta[6]	1.018	394.428	686.290
theta[7]	1.018	396.247	686.184
theta[8]	1.017	395.826	650.206
theta[9]	1.018	394.175	646.034
theta[10]	1.017	399.256	716.457
theta[11]	1.018	394.104	699.129
theta[12]	1.008	666.410	1096.062
theta[13]	1.018	392.954	680.476
theta[14]	1.017	395.068	661.907
theta[15]	1.017	397.589	678.013
theta[16]	1.018	393.037	675.383

Table 3. Diagnostics for the posterior estimates of σ.

variable	Rhat	ess_bulk	ess_tail
sigma	1.001	4965.037	2767.536

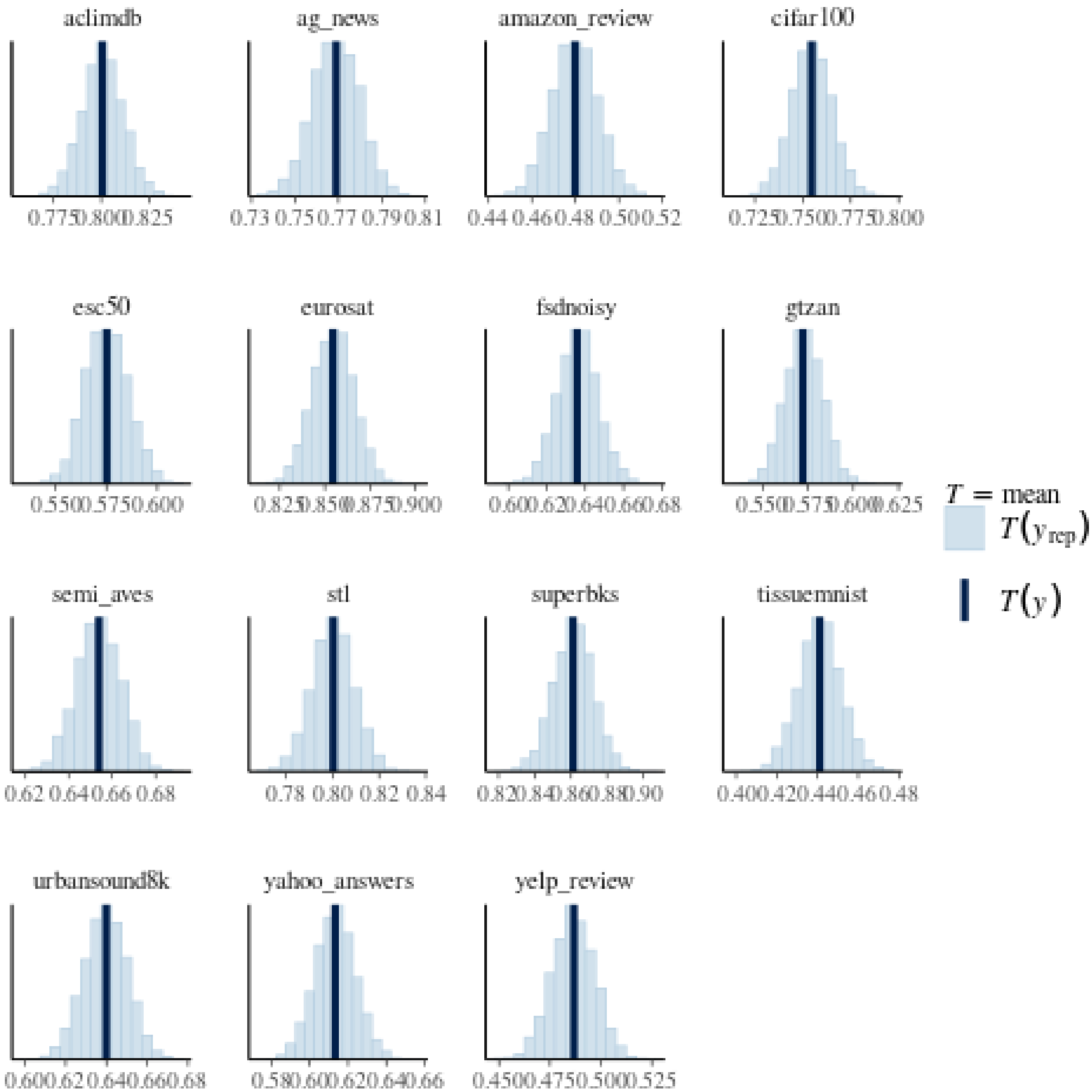

Figure 5. Posterior Predictive Checks

References

[1] Carlo A Furia, Robert Feldt, and Richard Torkar. Bayesian data analysis in empirical software engineering research. *IEEE Transactions on Software Engineering*, 47(9):1786–1810, 2019.

[2] Jonah Gabry, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. Visualization in bayesian workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182(2):389–402, 2019.

[3] David Issa Mattos, Jan Bosch, and Helena Holmström Olsson. Statistical models for the analysis of optimization algorithms with benchmark functions. *IEEE Transactions on Evolutionary Computation*, 25(6):1163–1177, 2021.

[4] Richard McElreath. *Statistical rethinking: A Bayesian course with examples in R and Stan.* CRC press, 2020.